CIRRELT

Centre interuniversitaire de recherche
sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre
on Enterprise Networks, Logistics and Transportation

An Empirical Comparison of Relocation Strategies in Real-Time Ambulance Fleet Management

Valérie Bélanger
Yannick Kergosien
Angel Ruiz
Patrick Soriano

December 2014

CIRRELT-2014-73

Document de travail également publié par la Faculté des sciences de l'administration de l'Université Laval, sous le numéro FSA-2014-012.

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Université Laval · McGill · Concordia · ÉTS · UQÀM · HEC Montréal · Polytechnique Montréal · Université de Montréal

An Empirical Comparison of Relocation Strategies in Real-Time Ambulance Fleet Management

Valérie Bélanger[1,2], Yannick Kergosien[3], Angel Ruiz[1,4], Patrick Soriano[1,2]

[1] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

[2] Department of Management Sciences, HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, Montréal, Canada H3T 2A7

[3] Université François-Rabelais de Tours, CNRS, LI EA 6300, OC ERL CNRS 6305, 64 av. Jean Portalis, Tours, 37200, France

[4] Department of Operations and Decision Systems, 2325 de la Terrasse, Université Laval, Québec, Canada G1V 0A6

Abstract. In order to ensure an adequate service to the population, Emergency Medical Services (EMS) rely on a given number of ambulances strategically located over the territory they serve. The arrival of calls to EMS being highly uncertain and evolving throughout the day, it may happen that at some point, the vehicles available to respond to these calls no longer cover properly all regions, even if the coverage was carefully planned initially. Relocation of ambulances may therefore be required during the day in order to achieve better performances. Some models tackling relocation have been proposed in the literature and it has been shown that using such strategies can help to improve overall performances. However, relocation generates movements that produce undesirable consequences from both economical and human resources management standpoints. Questions therefore arise: Is the relocation worth the effort? And if so, when and under which circumstances should it be used? What form should it take? Unfortunately, this issue has not been investigated much up to now. This study thus focuses on evaluating and analyzing relocation strategies, and reports extensive simulation experiments allowing to analyze the performance of these strategies when they face different levels of system workload.

Keywords. Emergency medical services (EMS), location, relocation, fleet management strategies, simulation.

Acknowledgements. This research was funded by the Fond de recherche du Québec - Nature et les technologies (FQRNT) through the Team research project program [grant PR-122269] and by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Discovery grants program [grants OPG 0293307 and OPG 0177174] and the Alexander Graham Bell Canada Graduate Scholarships-Doctoral Program awarded by the first author. We want to thank both organizations for their support.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

---

* Corresponding author: Valerie.Belanger@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2014

© Bélanger, Kergosien, Ruiz, Soriano and CIRRELT, 2014

1. Introduction

Emergency Medical Services (EMS) are a critical element of modern healthcare systems. In such systems, EMS are in charge of the pre-hospital component which consists of basic medical care and transportation activities performed from the reception of an emergency call to the release of the patient either in the care of a hospital or healthcare facility, or at the site of the incident if the patient does not require additional care. To provide this service to the population, EMS use a fleet of ambulances they locate strategically over the region they serve. However, even if the location of ambulances has been carefully planned in the initial static deployment/location plan, the uncertain nature of emergency calls, with respect both to their arrival times and their locations, may lead to a degradation of service level. Indeed, when some ambulances are dispatched to serve calls, the territory protected by the EMS will need to be covered with a reduced size fleet. Consequently, some areas of the region may be left without a proper coverage. Furthermore, the demand profile can also evolve significantly throughout the day such that a location plan adequate for a given time period may no longer be so for another. For these reasons, some corrective actions are generally required during the day in order to maintain or restore good performances, and one of the possible corrective actions is the relocation of ambulances.

Many studies dealing with ambulance fleet management show that the adoption of a relocation strategy can help in improving the system performance, at least for their specific context. Although it is clear that the relocation of ambulances represents an interesting means to maintain or restore a good service level, it also generates vehicle movements that lead to undesirable consequences from both economical and human resources management standpoints. Some questions therefore arise: Is relocation worth the effort? And if so, when and under which circumstances? What form should it take? Unfortunately, this issue has been rarely addressed in the literature. Indeed, Nair and Miller-Hooks [1] observed that many studies consider the development of relocation strategies, but none of them really seem to assess the benefits of using such strategies over more classical and static location ones. Nair and Miller-Hooks [1] therefore focus on the evaluation of a relocation strategy based on the results obtained solving a linear programming model. In this case, the use of linear programming requires a number of assumptions to be able to consider the different sources of randomness. Moreover, the strategy proposed by the authors is again only compared to the static case. In our opinion, even if it represents an interesting first step in the analysis of location and relocation strategies in a more realistic context, there is still a lot of work to do in this field.

The aim of this paper is to further the study of location and relocation strategies by performing a comparative analysis of different location and relocation strategies. This analysis seeks to quantify the benefits resulting from considering different dynamic strategies over more classical static ones. It also compares these strategies over various contexts with respect to various congestion levels. To conduct the analysis in a more realistic context, we have used a flexible and generic simulation model that was recently developed and presented in Kergosien et al. [2]. Indeed, simulation can help in dealing more adequately with different stochastic aspects inherent to EMS that cannot be addressed easily in the formulation of mathematical models. The contribution of this paper is therefore twofold. First, from a practical standpoint, by the formalizing and modelling of different location and relocation strategies. Secondly, by the detailed analysis of their respective benefits and drawbacks when confronted with classical static management strategies over varied realistic contexts.

The paper is organized as follows. A brief literature review on location and relocation strategies is presented in Section 2. In the third section, the different fleet management strategies considered in this study are described and the underlying mathematical models presented. Finally, in the last sections, the simulation model designed to analyze and assess each of these strategies is briefly presented, followed by the results of extensive simulation experiments and their analysis. A discussion on potential research avenues concludes the paper.

2. Location and relocation problems in EMS

Many researchers have studied different facets of EMS management. However, until recently, most of the work has essentially focused on the static location problem. The static location problem determines the set of standby sites where ambulances will be positioned while waiting to be dispatched to respond

to emergency calls. Once implemented, the corresponding location plan will remain unchanged, i.e. each ambulance will return to its designated standby site after completing a mission. Several approaches and models have been proposed in the literature to determine either the number and the location of ambulances required to achieve a given service level or, alternately, the location of a given ambulance fleet in order to maximize the system performance. Toregas et al. [3] were the first to explicitly formulate the ambulance location problem considering the widely used notion of coverage. Their model, the *Location Set Covering Problem*, seeks to minimize the number of vehicles needed to ensure that all demand zones can be reached (or covered) within a given time limit. The number of vehicles required to achieve such a coverage can however be very large and totally unrealistic in many practical contexts. In such situations, it would clearly be preferable to determine the best possible use for a vehicle fleet of a given size. Considering these practical limitations, Church and ReVelle [4] therefore formulate the *Maximal Covering Location Problem* which aims to maximize the population covered by a given vehicle fleet.

From these two seminal models, a significant number of models have been proposed to address the static location problem, among which we can cite the *Maximum Expected Covering Location Model* [5, 6], the *Maximum Availability Location Problem* [7], and the *Double Standard Model* [8] to name but a few. ReVelle [9], Marianov and ReVelle [10], Brotcorne et al. [11], Goldberg [12], Başar et al. [13], Bélanger et al. [14], and Aboueljinane et al. [15] present interesting surveys of the numerous models applied to emergency vehicle location, focusing mainly on the field of mathematical programming, but also on those of simulation and queueing theory. Since the focus of this work is on relocation strategies, we will not review further the literature on the static location problem and refer the interested reader to the excellent surveys listed above for details on the different models proposed to address it. The remainder of this section will instead be devoted to the description of the main relocation strategies with a particular attention to the objective pursued.

The static ambulance location mainly aims to select a set of standby sites where ambulances should be located between dispatches. Nevertheless, as discussed in the introduction, under certain circumstances, it may be beneficial to change ambulance locations or home bases during a day in order to take into account the evolution of the situation faced by the EMS. Doing so, one hopes to achieve a better service level to the population. Two main relocation strategies have been studied: *multi-period* and *dynamic* relocation strategies.

Multi-period relocation strategies consider that the demand pattern may fluctuate throughout the day due, among others, to population movements (e.g. early morning, morning commute, mid-day, evening commute, evening, night, etc.). A workday is then divided into several time periods, according to the various demand profiles, and different location plans are established a priori for each time period. At the end of a time period, ambulances will then move from their present location to the one they are assigned to for the next time period in order to implement that period location plan. Repede and Bernardo [16] formulate what is, to the best of our knowledge, the first multi-period ambulance location model. Their probabilistic model seeks to maximize the expected coverage while considering variations in both the demand pattern and the number of available ambulances. However, it does not explicitly account for relocation costs generated between periods. Carpentier [17] elaborates a deterministic model to consider the multi-period relocation problem. As opposed to the previous probabilistic model, it explicitly considers the movements between periods by integrating a term in the objective function to minimize relocations costs. This model thus minimizes costs involved in both serving emergency requests and relocating vehicles between periods given a fixed number of vehicles to locate at each period. Rajagopalan et al. [18] propose a probabilistic multi-period model that seeks the minimum number of vehicles needed to guarantee that each demand zone is covered with a given level of reliability. As is the case in Repede and Bernardo [16], this model does not take into account the relocation of vehicles between periods. Başar et al. [19] consider the problem of determining where and when ambulance stations (as opposed to standby sites) should be open over a multi-period planning horizon, such that a limited number of stations is used at each time period and that all the population is adequately covered by two distinct stations. This model differs from other multi-period models since it considers that when a station is open, it must remain so until the end of the planning horizon. Schmid and Doerner [20] present a multi-period model that considers travel time variations between periods, for instance due to road congestion. However, the number of vehicles and the demand patterns are assumed to be time-independent. The authors also propose to integrate in the objective function a penalty term to limit the number of relocated vehicle

between periods. Finally, Saydam et al. [21] extend the model proposed by Rajagopalan et al. [18] to consider the minimization of the number of vehicles that need to be relocated between periods in addition to the minimization of the number of vehicles needed to achieve an adequate coverage.

Multi-period location models presented here above take into account the fact that demand as well as the number of available vehicles and travel times may vary throughout a day. Although they are a better representation of what really happens in real-life, they do not explicitly consider the system state changes resulting from vehicle dispatches or ends of mission. To deal with these situations, one must resort to dynamic relocation strategies.

The first model that explicitly accounts for the dynamic nature of EMS when dealing with the relocation of ambulances is due to Gendreau et al. [22]. Their model determines the vehicle locations that maximize the population reachable by at least two vehicles within a given time limit, but simultaneously seeks to minimize relocations costs. In this case, relocation costs are directly related to the ambulances relocation history and aim to avoid long distance relocations, round-trips (i.e. relocations involving the same pair of sites), as well as to limit the movement of the same ambulance repeatedly. Gendreau et al. [23] present another dynamic relocation model, especially developed to address the location of physician cars (i.e. vehicles used to transport a physician to the scene of an emergency). The model determines how to locate a given physician car fleet such that the expected coverage is maximized. The problem is solved off line for each possible system state where the state of the system is defined as the number of cars available. Note that a limit on the number of vehicles that can be relocated between calls is imposed to prevent the fleet from being constantly relocated. Andersson and Värbrand [24] propose a dynamic ambulance relocation model that, rather than using a coverage measure as do most of the models presented so far, instead considers a measure called *preparedness* and defined as the capacity of the system to serve future demands. Relocation of vehicles is then launched in real-time when the preparedness level drops below a given threshold in order to regain a minimal preparedness level for each demand zone while minimizing the time required to achieve it. Similarly to Gendreau et al. [23], Nair and Miller-Hooks [1] formulate a multi-objective location-relocation model that considers the objectives of maximizing double coverage and minimizing location-relocation costs. As in the case of Gendreau et al. [23], the model is solved a priori for all possible system states. Maxwell et al. [25] consider dynamic programming to formulate an ambulance relocation problem seeking the policy that minimizes the discounted number of calls that cannot be reached in a timely manner. Schmid [26] also uses dynamic programming to address a relocation problem whose objective is to minimize the average response time over a finite planning horizon while considering travel time and demand density variations. In these two cases, relocation decisions are considered when a vehicle completes its mission and only the newly available vehicle is involved in the relocation process. Naoum-Sawaya and Elhedhli [27] elaborate a two-stage program to handle a priori relocation decisions. In this context, first stage decisions deal with the location of ambulances and second stage decisions with the assignment of emergency demands to ambulances. The objective of the model is to minimize the number of vehicles relocated as well as the number of demands that cannot be properly served. Finally, Mason [28] presents a dynamic ambulance relocation problem embedded into an EMS management software, that is, in many ways, similar to the one of Gendreau et al. [22]. Indeed, like the model of Gendreau et al. [22], this model aims to maximize service quality while minimizing relocation costs.

As stated earlier, the preceding literature review is not meant to be exhaustive, but covers the contributions we deemed to be the most relevant for this study. We now would like to highlight some observations that emerge from the analysis of these works. First of all, several models proposed to address both multi-period and dynamic relocation decisions integrate different mechanisms to limit relocation costs or inconvenience. Considering relocation costs in the development of such models is thus clearly a real and relevant issue. Moreover, it is possible to observe that a lot of effort has been deployed to develop relocation strategies, but most of them have only been compared with the static case and evaluated only in the specific context of the study in question (i.e. generally a particular EMS organization in a given environment). Further analyses of location and relocation strategies in realistic contexts that may differ in their congestion level or demand profile are clearly of interest. In addition, the comparison of different strategies having varied degrees of complexity against each other and not just with respect to the classical static case should help shed much needed light on their respective performances and costs as well as to under which conditions they seem more useful.

3. Fleet management strategies

This study focuses on the comparison and analysis of different management strategies related to the location and relocation of an ambulance fleet. Each strategy considered is characterized by a set of rules with respect to the initial location of vehicles, their repositioning and eventual relocation. More formally, the initial location of vehicles corresponds to the selection of the standby site for each vehicle at the beginning of their work shift. The repositioning consists in determining where to send a vehicle that has just completed a mission or resumes its service after a break. Finally, the relocation is concerned with the modification of a vehicle standby site over its work shift. Therefore, the combination of initial location, repositioning and relocation rules allow one to define different strategies. Here we propose to study and compare four strategies having varied levels of complexity. These strategies are summarized in Table 1 and described in this section. Note that for each strategy, we use as dispatching rule the one consisting in always assigning to a call the nearest vehicle.

Strategy	Initial location	Relocation	Repositioning
S1	A priori	-	Return to its initial location (fixed for the entire shift)
S2	A priori	Multi-period	Return to its initial location (with respect to the period)
S3	Dynamic		Dynamic
S4	Dynamic	Dynamic	Dynamic

Table 1: Management strategies studied

Each strategy presented in Table 1 results in a series of decisions, which are determined according to the objectives and constraints considered. Each strategy can therefore be represented by a distinct model. In the context of this study, all strategies have been formulated using the *Double Standard Model* (DSM) proposed by Gendreau et al. [8] as the reference model. Several reasons motivate this choice. First, the DSM is inspired by different governmental rules commonly used in practice, especially the *United States Emergency Medical Services (EMS) Act of 1973* [29]. The objective as well as the covering constraints it considers are easily understood and can be adapted to many cases with respect to location and relocation strategies. In fact, the DSM has already led to many variants and extensions including multi-period and dynamic models [22, 30, 20], and has been considered in several real-life situations [31]. Finally, formulating each strategy from the same model guarantees a comparison on common ground which is an essential attribute for this type of study. In this section, the DSM will be briefly described. We will then present the modifications that need to be introduced to adequately represent each of the proposed strategies in a context where the ambulance fleet size, the demand profile and the travel times can change according to the time period, and where work shifts vary and may overlap. Models formulated in this section will be used to replicate the decision process underlying each of the strategy within the simulation model.

The DSM integrates both concepts of double coverage and different coverage radii. It seeks to determine the location of a fixed number of vehicles in order to maximize the population covered twice within a defined time limit S given that at least a proportion α of the population of the region to serve must be covered within S and that all the population must be covered within S', with $S' > S$. This problem can be defined on a graph $G = (V \cup W, E)$ where $V = \{v_1, ..., v_n\}$ and $W = \{w_1, ..., w_m\}$ are two vertex sets representing, respectively, demand zones and potential standby sites, and $E = \{v_i, w_j\}$ is the edge set. To each edge (v_i, w_j) is associated a travel time t_{ij}. The population associated with demand zone $v_i \in V$ is equal to a_i. The number of ambulances is given and equal to p. The sets M_i and M_i' correspond to the sets of standby sites that can cover a demand zone v_i respectively within S and S'. The following variables and parameters are also used: x_j is an integer variable denoting the number of vehicles located at $w_j \in W$ with p_j an upper bound on x_j; y_i is a binary variable equal to 1 if and only if demand zone v_i is covered at least once within S; u_i is a binary variable equal to 1 if and only if demand zone v_i is covered at least twice within S. Therefore, the DSM can be formulated as follows:

DSM

$$\max \sum_{i=1}^{n} a_i u_i \tag{1}$$

subject to:

$$\sum_{j \in M_i'} x_j \geq 1, \ i = 1, ..., n, \tag{2}$$

$$\sum_{i=1}^{n} a_i y_i \geq \alpha \sum_{i=1}^{n} a_i, \tag{3}$$

$$\sum_{j \in M_i} x_j \geq u_i + y_i, \ i = 1, ..., n, \tag{4}$$

$$u_i \leq y_i, \ i = 1, ..., n, \tag{5}$$

$$\sum_{j=1}^{m} x_j = p, \tag{6}$$

$$x_j \leq p_j, \ j = 1, ..., m, \tag{7}$$

$$u_i, y_i \in \{0, 1\}, \ i = 1, ..., n, \tag{8}$$

$$x_j \geq 0, \ integer, \ j = 1, ..., m. \tag{9}$$

It is worth noting that, in certain cases, if the number of available vehicles is too small the model can become infeasible, i.e. all the population cannot be covered within S' and/or the proportion of the population that can be reached within S is less than α. It is also important to observe that the model indicates the number of vehicles to position at each of the potential standby sites, without any regard to the identity of the vehicles or paramedical teams that are located there. Indeed, the information on the number of vehicles to locate to each standby site is generally sufficient in the context of tactical planning, but can be inadequate when dealing with decisions of a more operational nature such as relocation. In the case of this study, it is necessary to know the exact location of each vehicle, in particular to quantify the impact of relocations (e.g. to compute relocation distances or times). Therefore, some modifications need to be made to the DSM to take these two aspects into account.

First of all, to ensure model feasibility at all times, constraints (2) and (3) will be dualized and their violations penalized in the objective function rather than considered as hard constraints. To do so, two new variables are introduced in the model to compute the constraints' violation: ρ_i, a binary variable equal to 1 if and only if demand zone v_i is not covered within S'; and δ, a variable equal to the population that cannot be reached by at least one vehicle within S. A second modification is also made to the objective function. Indeed, rather than considering directly the maximization of the population that is covered twice within S, we will instead maximize the probability that an emergency demand arises from a zone covered twice within S. By doing so, we believe that the model will represent better the actual need for an ambulance in a particular zone and how this need evolves over time as opposed to using the population of the zones which is a fixed data that does not evolve over the time of day. For instance, the need for ambulances in a residential sector will most certainly be different when considering the morning, before people leave for work, and the middle of the morning, after most working residents have left their homes. Given q_i, the probability that an emergency demand comes from demand zone v_i, and β_1, β_2 and β_3, weights associated respectively with the double coverage objective, the violation of the complete coverage constraint and the violation of the partial coverage constraint, the objective function of the DSM is modified as in (10). Furthermore, to take into account the identity of each vehicle in the decision process, variable x_j is replaced by variable x_{kj}, a binary variable equal to 1 if and only if vehicle k is located at w_j. A modified version of the Double Standard Model (mDSM) can now be formulated as follows:

mDSM

$$\max \beta_1 \sum_{i=1}^{n} q_i u_i - \beta_2 \sum_{i=1}^{n} \rho_i - \beta_3 \delta \tag{10}$$

subject to:

$$\sum_{j \in M_i'} \sum_{k=1}^{p} x_{kj} \geq 1 - \rho_i, \ i = 1, ..., n, \tag{11}$$

$$\sum_{i=1}^{n} a_i y_i \geq \alpha \sum_{i=1}^{n} a_i - \delta, \tag{12}$$

$$\sum_{j \in M_i} \sum_{k=1}^{p} x_{kj} \geq u_i + y_i, \ i = 1, ..., n, \tag{13}$$

$$u_i \leq y_i, \ i = 1, ..., n, \tag{14}$$

$$\sum_{j=1}^{m} x_{kj} = 1, \ k = 1, ..., p, \tag{15}$$

$$\sum_{k=1}^{p} x_{kj} \leq p_j, \ j = 1, ..., m, \tag{16}$$

$$u_i, y_i, \rho_i \in \{0, 1\}, \ i = 1, ..., n, \tag{17}$$

$$x_{kj} \in \{0, 1\}, \ j = 1, ..., m, \ , k = 1, ..., p, \tag{18}$$

$$\delta \geq 0. \tag{19}$$

The mDSM will thus serve as a basis in the modelling of the different location and relocation strategies studied hereafter.

3.1. Strategy 1: A priori initial location without relocation.

The first strategy consists exclusively of a set of static rules. In this case, the initial location of a vehicle is determined a priori and fixed for the entire work shift. The vehicle is sent to its initial standby site at the beginning of its shift and will be sent back there every time it completes a mission or resumes its service. Therefore, vehicle availability as well as beginnings and ends of work shifts need to be taken into account in the corresponding model. Consequently, this model will differ from the mDSM since it has to consider the multi-period aspect of the problem to properly integrate the information regarding the fleet availability over a day. We will therefore assume that the problem is formulated within a planning horizon composed of time periods and that the different time periods correspond to the changes in the conditions faced by the EMS over time, e.g. changes in the probability that a demand arises from a particular zone, changes in travel times, etc. We will also assume that time periods are defined so that the beginning and the end of work shifts coincide with them. Nevertheless, as opposed to most of the multi-period models proposed to address the ambulance location problem, in this static strategy, relocation is not allowed between periods: once the standby site is determined for a vehicle, it will remain the same until the end of its work shift.

To adequately represent this specific context, the mDSM needs to be adapted. Clearly, all parameters and variables defined previously need to be modified in order to integrate a time period superscript t, with the exception of a_i and p_j that will be considered as time-independent. For instance, variable u_i becomes u_i^t, a binary variable equal to 1 if and only if a demand zone v_i is covered twice within S during period t. Three new parameters also need to be introduced to account for vehicle availability: d_k^t, a parameter equal to 1 if vehicle k is available at t, $t = 1, ..., T$, and 0 otherwise; t_b^k, the time period corresponding to the beginning of vehicle k work shift; and t_e^k, the period corresponding to the end of vehicle k work shift. Taking into account these new variables and parameters, the Multi-Period Double Standard Model (MPDSM) is formulated as follows:

MPDSM

$$\max \sum_{t=1}^{T} (\beta_1 \sum_{i=1}^{n} q_i^t u_i^t - \beta_2 \sum_{i=1}^{n} \rho_i^t - \beta_3 \delta^t) \tag{20}$$

subject to:

$$\sum_{j \in M_i^t} \sum_{k=1}^{p} x_{kj}^t \geq 1 - \rho_i^t, \ i = 1, ..., n, \ t = 1, ..., T, \tag{21}$$

$$\sum_{i=1}^{n} a_i y_i^t \geq \alpha \sum_{i=1}^{n} a_i - \delta^t, \ t = 1, ..., T, \tag{22}$$

$$\sum_{j \in M_i^t} \sum_{k=1}^{p} x_{kj}^t \geq u_i^t + y_i^t, \ i = 1, ..., n, \ t = 1, ..., T, \tag{23}$$

$$u_i^t \leq y_i^t, \ i = 1, ..., n, \ t = 1, ..., T, \tag{24}$$

$$\sum_{j=1}^{m} x_{kj}^t = d_k^t, \ k = 1, ..., p, \ t = 1, ..., T, \tag{25}$$

$$\sum_{k=1}^{p} x_{kj}^t \leq p_j, \ j = 1, ..., m, \ t = 1, ..., T, \tag{26}$$

$$x_{kj}^t \leq x_{kj}^{t+1}, \ j = 1, ..., m, \forall k : t_b^k < t_e^k, t = t_b^k, ..., t_e^k - 1, \tag{27}$$

$$x_{kj}^t \leq x_{kj}^{t+1}, \ j = 1, ..., m, \forall k : t_b^k > t_e^k, \ t = 0, ..., t_e^k - 1 \ , \ t = t_b^k, ..., T - 1, \tag{28}$$

$$x_{kj}^T \leq x_{kj}^0, \ j = 1, ..., m, \forall k : t_b^k > t_e^k, \tag{29}$$

$$u_i^t, y_i^t, \rho_i^t \in \{0, 1\}, \ i = 1, ..., n, \ t = 1, ..., T, \tag{30}$$

$$x_{kj}^t \in \{0, 1\}, \ j = 1, ..., m, \ , k = 1, ..., p, \ t = 1, ..., T, \tag{31}$$

$$\delta^t \geq 0, \ t = 1, ..., T. \tag{32}$$

The MPDSM thus integrates several constraints that aim to ensure that a vehicle is located at some standby site only during its service (25), and that once its standby site is determined, it remains the same for its entire work shift (constraints (27) to (29)). The MPDSM shares similarities with the multi-period model presented by Başar et al. [19] since it considers that when a location is open, it must remain so until the end of the planning horizon. This model also considers coverage objectives similar to the ones pursued by the MPDSM. However, the problem studied in [19] deals with a strategic location problem concerned with building fixed stations as opposed to here, where one considers instead the location of an ambulance fleet over a short planning horizon.

3.2. Strategy 2 : A priori initial location with multi-period relocation

The second strategy we propose to study still considers that the initial location of vehicles is determined a priori but relocation of vehicles between periods is now allowed. A location plan is then determined for each time period and the standby site assigned to a given vehicle can be changed from one period to the next. After completing a mission or resuming its service, the vehicle will be sent back to the standby site it is assigned to in the current period. The corresponding model is thus quite similar to the one of the first strategy. The MPDSM will therefore need only minor modifications in order to model the second proposed strategy. Indeed, since vehicle relocation between periods is now allowed, the constraints that fix standby sites at the beginning of work shifts (i.e. constraints (27) to (29)) are no longer needed The resulting model could then be decomposed into T independent models, one for each time period. However, doing so does not allow to explicitly account for relocation costs or inconvenience.

To properly consider relocation inconvenience we propose the following Multi-Period Double Standard Model with Relocation (MPDSMR) which is based on the MPDSM. As mentioned earlier, the MPDSMR will differ from the MPDSM since constraints (27) to (29) have been removed. In addition, the MRMPDS will consider two objectives: a primary objective that still aims to maximize system performance and the satisfaction of coverage constraints (as in the MPDSM objective function) and a secondary objective that seeks to minimize relocation costs. Given, $r_{kjj'}^t$, a binary variable equal to 1 if and only if vehicle

k is moved from standby site w_j to standby site $w_{j'}$ at the end of period t, and $c^t_{kjj'}$, relocations costs generated when vehicle k is moved from standby site w_j to standby site $w_{j'}$ at the end of period t, the secondary objective is formulated as follows:

$$\min \sum_{t=1}^{T} \sum_{k=1}^{p} \sum_{j=1}^{m} \sum_{j'=1}^{m} c^t_{kjj'} r^t_{kjj'} \tag{33}$$

The following constraints also need to be included in the model to adequately compute relocation costs incurred between periods:

$$r^t_{kjj'} \geq x^t_{kj} + x^{t+1}_{kj} - 1, \; j = 1, ..., m, \; j' = 1, ..., m, \; k = 1, ..., p, \; t = 0, ..., T-1, \tag{34}$$

$$r^T_{kjj'} \geq x^T_{kj} + x^0_{kj} - 1, \; j = 1, ..., m, \; j' = 1, ..., m, \; k = 1, ..., p. \tag{35}$$

The MPDSMR is therefore similar to the model proposed by Schmid and Doerner [20]. Both models allow vehicles to be relocated between periods and consider the associated cost as an objective to minimize. However, MPDSMR considers vehicle availability over a day more explicitly.

3.3. Strategy 3 : Dynamic initial location and repositioning.

The third strategy that we propose to analyze considers that decisions related to the initial location of vehicles and their repositioning are taken in real time rather than *a priori*. The initial location of a vehicle is then determined with respect to the system state, in particular the location of on-duty and available vehicles, at the exact moment the vehicle starts its work shift. The standby site where a vehicle will be sent at the end of a mission or where it resumes its service will also be selected in real-time, meaning that a vehicle can be assigned to different standby sites over the duration of its work shift. However, no relocation of vehicle waiting to be dispatched to emergency calls is allowed as in [25, 26].

To follow the same logic as the models presented previously, the standby site of a vehicle at the beginning of its work shift, at the end of a mission, and when it resumes its service after a break, is selected considering two criteria. First, if all demand zones cannot be covered within S', the newly available vehicle will be sent to the standby site that maximizes the number of zones covered within S' and whose capacity is not saturated. Thus, the standby site is chosen in order to reduce the problem infeasibility with respect to the primary objective of the double standard model principle. If several equivalent solutions are available from this feasibility perspective, the one corresponding to the standby site that maximizes system performance will be selected, i.e. the standby site that maximizes the number of zones covered twice within S weighted by the probability that a demand arises in the zone. Finally, if all demand zones can be covered within S', the newly available vehicle will be sent to the standby site that maximizes the number of zones covered twice within S weighted by the probability that a demand arises in the zone and whose capacity is not full. In this last situation, only the maximization of the system performance is considered.

3.4. Strategy 4 : Dynamic initial location, repositioning and relocation.

The last strategy consists of the whole set of dynamic rules. As opposed to the previous strategies, the dynamic relocation of available vehicles in real-time is now permitted. The location of vehicles waiting to be dispatched to emergency calls can therefore be modified if the system state requires it. Here, we consider that the system performance justifies such a relocation procedure when all demand zones are no longer reachable within S'. Clearly, if the number of available vehicles during a portion of the day is small, relocation procedures could be launched very frequently, which would be counter productive since vehicles would be constantly moving around as well as difficult to manage from a human resources perspective. Hence, to prevent such unwanted side effects, relocation decisions will be allowed only if a predefined amount of time has elapsed since the last relocation occurred. We will refer to this minimal time between two consecutive relocation procedures as τ. Moreover, the relocation of a specific vehicle will be penalized if that vehicle has been moved more recently than that threshold, i.e. its last repositioning occurred less than τ units of time ago.

Two main events can lead to relocation decisions: 1) the appearance of a vehicle after completing a mission or resuming its service after a break; 2) the disappearance of a vehicle from the system following

either its dispatch to a call, the beginning of a break or at the end of its work shift. In the first situation, if all demand zones cannot be served within S' and the last relocation process occurred more than τ units of time ago, a relocation is launched, including the repositioning of the vehicle that just appeared in the system. If no relocation is required nor allowed, the appearing vehicle standby site is selected according to the rule described in *Strategy 3*. In the second situation, if the complete coverage within S' cannot be maintained and the last relocation occurred more than τ units of time ago, a relocation procedure is launched. Otherwise, all vehicles stay at their current standby sites.

To adequately model the dynamic relocation problem described here above, the following variables and parameters need to be introduced: $\hat{P}$ is the set of vehicles whose relocation generate a penalty β_4 (i.e. those for which the last repositioning occurred less than τ units of time ago); λ_{kj} is a parameter equal to 1 if vehicle k is located at w_j before the relocation process and 0 otherwise; and s_k is a binary variable equal to 1 if and only if vehicle k is relocated (i.e. its standby site is changed by the relocation procedure). Therefore, the Double Standard Model with Dynamic Relocation considered at time t (DSMDR$_t$) can be formulated as follows:

DSMDR$_t$

$$\max \beta_1 \sum_{i=1}^{n} q_i u_i - \beta_2 \sum_{i=1}^{n} \rho_i - \beta_3 \delta - \beta_4 \sum_{k \in \hat{P}} s_k \tag{36}$$

$$\min \sum_{k=1}^{p} \sum_{j'=1}^{m} \sum_{j=1}^{m} \lambda_{kj'} c_{kj'j} y_{kj} \tag{37}$$

subject to:

$$\sum_{j \in M_i'} \sum_{k=1}^{p} x_{kj} \geq 1 - \rho_i, \ i = 1, ..., n, \tag{38}$$

$$\sum_{i=1}^{n} a_i y_i \geq \alpha \sum_{i=1}^{n} a_i - \delta, \tag{39}$$

$$\sum_{j \in M_i} \sum_{k=1}^{p} x_{kj} \geq u_i + y_i, \ i = 1, ..., n, \tag{40}$$

$$u_i \leq y_i, \ i = 1, ..., n, \tag{41}$$

$$\sum_{j=1}^{m} x_{kj} = 1, \ k = 1, ..., p, \tag{42}$$

$$\sum_{k=1}^{p} x_{kj} \leq p_j, \ j = 1, ..., m, \tag{43}$$

$$\sum_{j=1}^{m} \lambda_{kj} x_{kj} - s_k \geq 0, \ k = 1, ..., p, \tag{44}$$

$$u_i, y_i, \rho_i, \in \{0, 1\}, \ i = 1, ..., n, \tag{45}$$

$$x_{kj} \in \{0, 1\}, \ j = 1, ..., m, k = 1, ..., p, \tag{46}$$

$$s_k \in \{0, 1\}, \ k = 1, ..., p, \tag{47}$$

$$\delta \geq 0. \tag{48}$$

The DSMDR$_t$ is therefore a multi-objective model that aims to, firstly, maximize the double coverage while penalizing the violation of constraints (38) and (39) as well as the relocation of recently moved vehicles, and secondly, minimize relocation costs. Constraints (44) check if a given vehicle has been relocated or not. Other constraints are essentially equivalent to those of the previous models presented.

Obviously, several criteria can be envisioned to properly implement dynamic relocation strategies in real-life applications, i.e. when to launch relocation, which vehicles to relocate, etc. In the context of this

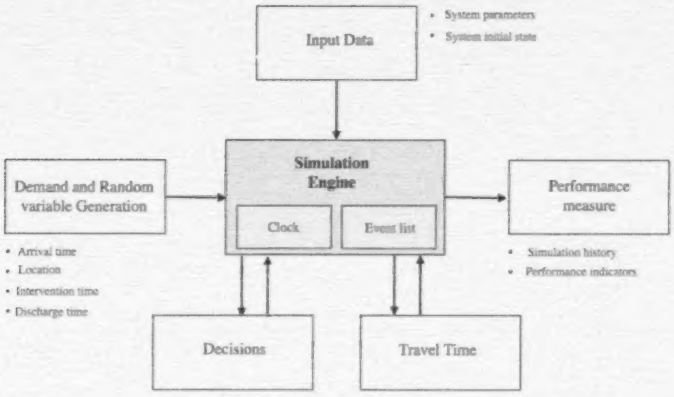

Figure 1: Simulation model architecture

study, dynamic relocation has been implemented in order to limit both the number and the frequency of relocations by imposing a minimum elapsed time τ between two successive relocation procedures. However, this limit should be set in a way that the number of relocations is large enough to really assess the impact of such a strategy.

The four strategies presented here and their underlying models will be implemented in the simulation model developed to analyze EMS management. They will then be compared on the basis of their performances over several simulation experiments.

4. Simulation model

To compare and analyze the different fleet management strategies, a flexible and generic object-oriented simulation model has been developed. Indeed, a simulation approach allows the consideration of the various sources of uncertainty inherent to EMS management. The proposed model is built on the following components: an *Input Data* block that includes all the parameters needed to an adequate description of the system and its initial state, a *Demand and Random Variable Generation* block that generates all emergency demands as well as all random variables required to conduct the simulation study, a *Simulation Engine* that manages the simulation itself, i.e. the simulation clock and the list of events, a *Travel Time* block that estimates ambulance travel times between two locations of the region to serve, a *Decisions* block that manages the decision making process required to adequately replicate EMS operations at the appropriate moment during the simulation, and finally, a *Performance Measure* block that traces and compiles all the information required to evaluate the system performance. The model architecture is illustrated in Figure 1 and each of its elements is now briefly presented. Note that this simulation tool has been developed in a more general context and has been designed to be able to analyze other types of decisions at all decision-making levels. We refer the reader to Kergosien et al. [2] for a more detail description of the simulation tool, from its design to its validation.

4.1. Input Data

This block contains all the information allowing a proper description of the system and its initial state. Input data is classified into two categories: data related to system characteristics and data specifying the system's initial state including the initial location of ambulances and the initial state of each resource among others. System characteristics define the resources available and the region under the EMS responsibility. It describes the territory to serve, its division into zones and districts, the list of available ambulances, the set of potential standby sites, and the set of hospitals or health facilities in the region.

4.2. Demand and Random Variable Generation

An important part of computer simulation consists of sampling probability distributions in order to draw plausible specific values for uncertain events to be used during the simulation execution. For example, in the context of an EMS, service time at an emergency site is, in practice, unknown at the time the emergency call is received. Service times can however be modelled by a probability distribution function of known parameters, which can be determined from historical data for instance. The specific value for the service time of a given intervention can be determined a priori or during the simulation experiment from the associated probability distribution. In the simulation tool used here, all random variables are generated a priori and stored in an external file. These variables are then used at the appropriate time during a simulation experiment. This technique has been selected in order to allow a better control of the model and to increase its flexibility. Moreover, it eases the validation and the verification process. Finally and more importantly, generating as much random events as possible out of the simulation execution drastically reduces the variance of the simulation results.

4.3. Simulation Engine

The model used in this paper is based on Discrete Event Simulation (DES). DES, as defined in Law [32], models a system as it evolves over time. The system thus changes instantaneously at particular points in time corresponding to events that modify the system state. The system is defined by a set of entities each characterized by a given number of attributes and variables that may evolve over time. In our case, the model is defined by three main types of entities: emergency requests, ambulances and operators (including emergency medical respondents and dispatchers). The management of events is provided by a *simulation engine*, a timing routine which is in charge of moving the simulation clock from one event time to the following. The simulation engine of the tool is inspired by the one proposed by Pidd [33], which consists of a three-phase algorithm that allows the clock to be advanced asynchronously from one event to the next. The simulation engine works with a list of events sorted by their execution time. Each time an event occurs, some decision procedures are triggered and the state of the system and its entities are modified or adjusted according to the decisions taken. The simulation process then resumes moving the simulation clock to the next event.

4.4. Travel Time

Simulating travel times is a very difficult yet important task. In fact, travel times influence the precision of the simulation results and are a key element when evaluating the credibility of the simulator. Several methods can be used to estimate travel times, including sophisticated methods linked to powerful geographic information systems (GIS). In the simulation tool used here, we chose to implement the following relatively simple and generic method which is based on a priori knowledge of some real travel times or estimates for a set of important or frequent locations (e.g. hospitals, potential waiting sites, zone centroids). Let M be the matrix of known travel times between these locations, the size of M depends on the amount of information that can be obtained from the real case studied. Evidently, increasing the number of points in M will increase the accuracy of the estimated travel times. During the simulation, the computation of travel time t_{ab} between two locations a and b not in matrix M, is based on the known travel time $t_{a'b'}$ between two locations a' and b', where a' and b' are the locations in M that are the nearest to a and b respectively, and on the Euclidean distances from a to b and from a' to b', noted d_{ab} and $d_{a'b'}$, as follows : $t_{ab} = (d_{ab} * t_{a'b'})/d_{a'b'}$. Clearly, this method is not as accurate as a GIS based one. However, if matrix M contains enough points, this method should approximate adequately travel times by taking into account through the data in the matrix the presence of obstacles or particular features of the transportation infrastructure (e.g. highways, bridges, tunnels, one-ways) as well as the general traffic conditions on the itineraries corresponding to each pair of locations in M.

4.5. Decisions

The execution of some events during the simulation experiments implies the modelling and the replication of some decision processes carried out by operators or dispatchers from EMS organizations. In particular, the *Decisions* block implements the repositioning and relocation decisions as described in the previous section.

4.6. Performance Measure

The last block is devoted to support the analyses of the simulation results by the user. To this end, a complete history of the simulation is used to calculate some performance indicators. This history includes all movements performed by each ambulance, all the times at which the entities and resources states changed, and other statistical information about the decisions taken. Clearly, the performance measures to compute depend on the particular goals of the study.

4.7. Implementation, verification and validation

The simulation model was implemented in C++. The use of a generic programming language was justified by the need of higher flexibility and to avoid restriction due to specific architectures of simulation softwares. Moreover, this approach allows to easily build routines that will replicate almost any decision procedure. The different models proposed to represent each of the management strategies considered in the *Decisions* block are solved with *CPLEX 12.5*, except for *Strategy 3* where the best possible repositioning decision is determined by complete enumeration of all the possibilities. For more details about the verification and the validation process, we refer the reader to Kergosien et al. [2].

5. Simulation experiments

Extensive simulation experiments were carried out to evaluate and compare the performance of the four fleet management strategies presented in section 3. To this end, we generated a set of 50 random instances based on a fixed demand profile. These instances correspond to 50 realizations of demands and all random variables associated to them, which were obtained through replications generated by the Demand and Random Variable Generation block of the simulation tool. Each instance was then treated using the four proposed strategies S1 to S4, and for each case we replicated the strategies for four different fleet sizes. Overall, 800 executions were thus completed in order to draw a good appraisal of the strategies' behaviours to different levels of congestion.

The characteristics of the region on which the simulations were performed are based on the topology and population distribution of the Montreal region, the major population center in the province of Québec (Canada). However, the demand and service time data were generated as detailed hereafter. In addition, the strategies and rules such as location and relocation policies as well as dispatching decisions may not correspond to the ones actually used by the local EMS of Montreal (Urgences-santé[1]) for which we have no official information. In our study, location and relocation strategies correspond to the ones proposed in section 3. Moreover, recall that the nearest ambulance is always dispatched to a call. Finally, we assume that the fleet is managed in a centralized manner such that a sole decision maker manages the whole region and all the ambulances.

The region considered in the study contains 600 zones, 40 potential standby sites arbitrarily located, two depots, and 15 hospitals or health facilities. A total of 125 paramedical teams working on 8-hour shifts are considered in the base case. The number of teams on duty at each period of the day changes according to a curve of demand inspired from Ingolfsson et al. [34].

Since no real data was available to us, we randomly generated a set of realistic data by merging several sources of information: annual reports of the local EMS organization [35], population statistics for the region [36], and information collected from the literature. As mentioned in Kergosien et al. [2], most of these informations being of an aggregated nature, we therefore set the parameters of our generator empirically in order to ensure that the detailed data generated was effectively realistic and adequately fitted the aggregate data collected in terms of total number of requests, number of teams and size of the fleet.

As is generally accepted in the literature [37], an exponential distribution was used to model the inter-arrival times between two consecutive emergency demands. Each day was divided into 12 two-hour periods to build a daily workload curve which accounts for the variation in demand intensity throughout the day. The mean of the exponential distribution was arbitrarily set to a specific value for each of these periods, ranging from 1.5 to 5 minutes. Once a emergency demand is generated, it is associated to a

[1] https://www.urgences-sante.qc.ca/

specific zone following a discrete distribution where the probability of selecting a zone depends on its demographic weight. As discussed in Aboueljinane et al. [15], this is one of the approaches proposed to adequately generate demands.

The intervention time at the scene of emergency as well as the discharge time at the hospital including the transfer of the patient and some administrative tasks, have been modelled using Gamma(k, θ) distributions In 75 % of the cases, emergency demands will need the patient to be transported to a hospital. If such is the case, the intervention time then follows a Gamma(3, 5) distribution and the discharge time, a Gamma(8, 5) distribution. The hospital or health facility to which the patient will be transported is selected randomly with a strong bias towards the closest one. When no transportation is required, the intervention time follows a Gamma(3,10) distribution. The choice of these distributions and parameters is in line with those used in Schmid [26]. Finally, travel time from a particular location to another is computed as described at the end of section 4.4 using the Euclidean distance between each point as well as a priori knowledge of some real travel times or estimates for a set of important locations. In all cases, the values of S and S' are set to 9 and 11 minutes respectively. When dynamic relocation is considered, the value of τ is set to 15 minutes.

We refer to the 50 instances generated using the parameter and distributions presented here above as the base case (B). However, to compare location and relocation strategies in varied contexts, the base case was treated considering different fleet sizes. In particular, we reduced by 20 % and 10 % the fleet size in experiments RV20 and RV10, respectively, and we increased it by 10 % in experiment IV10. Thus, four series of experiments encompassing fleets of 101, 113, 125, and 138 vehicles named RV20, RV10, B, and IV10, respectively, were conducted on the 50 instances, for a total of 800 executions. Note that the increase/reduction affects the number of available vehicles at each time period homogeneously.

To analyze and compare the location and relocation strategies, two types of performance measures were computed and reported in Table 2. The first type is related to service performance and includes average response time, noted RT, and the percentage of calls reached within 9 minutes, %≤9. The second type concerns fleet management efficiency and reports total traveled distance, TTD, and relocations costs. In order to adequately express what relocation costs may represent for the organization, we computed two measures, total relocation distance, TRD, and the total number of relocations performed, #R. Each instance spans a period of 7 consecutive simulated days but only the 5 middle are used to compute performance measures. Proceeding this way removes the transient states corresponding to the beginning and the end of the simulation. Complete numerical results, including the half-width of the confidence intervals at 95 % for each performance measure are reported in Table 2. As can be observed, the values of the confidence intervals are generally quite low, which confirms the statistical precision of the computed estimates.

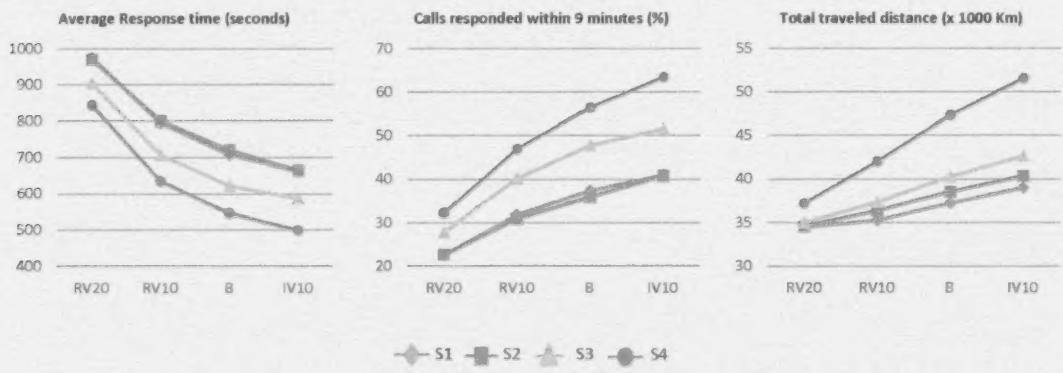

Figure 2: Results for variable ambulance fleet sizes

To better illustrate the information reported in Table 2, Figure 2 depicts the performance produced by strategies S1 to S4 to the four sets of instances. We will use both in our discussion. Figure 2 confirms

		S1			S2			S3			S4		
RV20	RT	973	±	10	969	±	11	903	±	11	843	±	12
	%≤9	22.5	±	0.4	25.0	±	0.4	27.8	±	0.5	32.1	±	0.6
	TTD	34 449	±	171	34 535	±	154	34 961	±	159	37 162	±	184
	TRD	0.0	±	0.0	177	±	20	0.0	±	0.0	4913	±	112
	#R	0.0	±	0.0	102.9	±	5	0.0	±	0.0	838.4	±	18
RV10	RT	794	±	6	800	±	6	707	±	7	634	±	7
	%≤9	31.7	±	0.3	30.8	±	0.4	40.2	±	0.5	46.8	±	0.5
	TTD	35 232	±	153	36 355	±	147	37 347	±	130	41 961	±	136
	TRD	0.0	±	0.0	409	±	26	0.0	±	0.0	7169	±	112
	#R	0.0	±	0.0	209.6	±	7.3	0.0	±	0.0	1 232.0	±	20.5
B	RT	707	±	5	720	±	4	622	±	5	545	±	4
	%≤9	37.3	±	0.4	35.9	±	0.3	47.7	±	0.5	56.4	±	0.5
	TTD	37 217	±	142	38 525	±	134	40 273	±	158	47 285	±	143
	TRD	0.0	±	0.0	553	±	32	0.0	±	0.0	8 101	±	84
	#R	0.0	±	0.0	378.5	±	9.7	0.0	±	0.0	1 391.8	±	14.0
IV10	RT	661	±	3	664	±	3	586	±	3	498	±	2
	%≤9	40.8	±	0.3	40.7	±	0.4	51.5	±	0.4	63.3	±	0.3
	TTD	38 967	±	152	40 301	±	151	42 617	±	156	51 509	±	153
	TRD	0.0	±	0.0	893	±	37	0.0	±	0.0	7 985	±	76
	#R	0.0	±	0.0	566.7	±	9.5	0.0	±	0.0	1 355.9	±	11.5

Table 2: Complete numerical results

two observations that apply to the four strategies considered. First, the service performance increases in a non-linear manner as the size of the fleet increases. Second, the total traveled distance increases as the size of the fleet increases. We will discuss this behaviour later on.

A more thorough look at Table 2 and Figure 2 allows also to conclude that there is no significant difference between strategies S1 and S2 with respect to the service performance. Moreover, Table 2 confirms that confidence intervals produced by S1 and S2 for RT overlap in three out of four cases and, those produced for %≤9 overlap in one case. We also observe that S2 leads to larger traveled distances than S1. Therefore, although many studies claim that multi-period relocation helps improving service level, our empirical conclusion is that S2 is dominated by S1. This counter-intuitive conclusion can be explained, at least partially, by the way S1 is defined in this study. Indeed, recall that in S1, the initial location of a vehicle is determined a priori and fixed for its entire work shift. Also, vehicles return to their assigned standby location every time a mission is completed. Therefore, S2 seems to have an important advantage with respect to S1 because relocations can be made at each time period. However, since ambulance shifts are managed in a flexible manner conceived to fit the demand profile, S1 chooses at each period the standby location of the ambulances starting their shift in such a way that the total coverage is optimized. In other words, S1 has the opportunity to adapt partially the fleet deployment according to the evolving demand, through the arrival of ambulances which start their working shift.

Unsurprisingly, Figure 2 shows that dynamic strategies S3 and S4 produce better service performances than S1, but at the expense of higher total traveled distances. More specifically, when compared to S1, dynamic repositioning (S3) is able to reduce significantly the average response time by 70 to 88 seconds, depending on the group of instances considered, as well as increase the percentage of calls that can be reached within 9 minutes by 5 to 11 %. Adapting the system one vehicle at a time can thus clearly lead to interesting service performance improvements. On the other hand, and always with respect to S1, average total traveled distance is increased by 1.5% in instances RV20, and up to 9.4% in the case of IV10.

Finally, S4, which performs both dynamic repositioning and dynamic relocations, leads to the best service performance. When compared to S1 it reduces the average response time by 130 to 162 seconds, which represents as much as twice the improvement achieved by S3. The average percentage of calls reached within 9 minutes ranges from 32.09 to 63.28%, which means improvements of 9.59% up to 22.48% with respect to the ones produced by S1. However, to be able to get these improvements, S4 performs on average between 838 and 1391 relocations, and increases total traveled distance by 7.88% up

to 31.74% with respect to S1.

It is worth mentioning how system saturation or congestion impacts the behaviour shown by the more flexible strategies S2 and S4 in terms of relocation costs. Indeed, both the number of relocations performed by S2 and S4 as well as the total traveled distance increase when more vehicles are available. In other words, when the capacity of the fleet is tighter with respect to the demand as in RV20, vehicles are busy almost all the time and there is not much opportunity to perform relocations. As the number of vehicles increases, some slack is created in the system and thus the number of relocations increases. The total traveled distance also increases because as calls arrive and the vehicles that should answer are already busy, other vehicles located further away must respond to the calls.

The results presented in Figure 2 also allow to conclude that using less vehicles with dynamic location and relocation strategies can allow to achieve similar performances than increasing the ambulance fleet size. Indeed, S3 and S4 lead to better results than those obtained by increasing the base case (B) fleet size by 10 % and using S1. Moreover, reducing the fleet size by 10% and using dynamic strategies leads to results equivalent than the ones obtained on the base case considering static strategies. Therefore, dynamic strategies can be considered as very interesting alternatives over the increase the number of ambulances. However, when the system is really congested, even if dynamic strategies can help in improving the system, it cannot compensate for the lack of resources.

To summarize, our numerical study clearly shows that a better service performance can be achieved by using more flexible management strategies. It also quantifies the additional cost (both in terms of total traveled distance and number of relocations) required by the additional flexibility. Finally, it suggests that, when the capacity of the fleet is tight, flexible management strategies are the only way to achieve reasonable coverage. Therefore, dynamic strategies can theoretically be considered as very interesting alternatives when resources cannot be increased. In our opinion, and based on the numerical results presented herein, repositioning strategy such as that defined by S3 might be a good option to envision for those organizations currently managing their fleet by a static deployment strategy. Relocation within certain limits might also be really interesting. The better strategy to use will however always depend on the context under study and the organization involved in the process.

6. Conclusion

This paper presents and formalizes location and relocation strategies to manage ambulance fleets. More specifically, four management strategies have been proposed and modelled considering different initial location, repositioning and relocation strategies, ranging from simple strategies to sophisticated ones. Each strategy corresponds to series of decisions which lead to a specific mathematical model. All strategies have been modelled according to the *Double Standard Model* [8] to ensure a fair comparison of their performances. A simulation tool has been used to perform extensive numerical experiments on random instances inspired by the context of a major canadian city. The main objective of these experiments was to quantify the benefits of considering dynamic strategies over more classical static ones, but also to compare those strategies in various contexts with respect to their congestion level. The experiments show that dynamic location and relocation strategies clearly lead to better service performances which can be seen as equivalent to an increase in the size of the ambulance fleet. However, dynamic strategies also generate significant increases in the total traveled distance and eventually, the number of relocations, which might be difficult to justify for some organizations, in particular with respect to human resources management.

The development and analysis of tactics and strategies that will allow the limitation or control the costs related to systematic dynamic relocations, leading to hybrid policies between dynamic repositioning and full dynamic relocations, could be of high interest. In this way, relocation of ambulances would be considered, but in a controlled manner both in terms of frequency and intensity. The current analysis should be extended to other case studies in order to validate it on different contexts.

References

[1] R. Nair, E. Miller-Hooks, Evaluation of relocation strategies for emergency medical service vehicles, Journal of the Transportation Research Board 2137 (2009) 63–73.

[2] Y. Kergosien, V. Bélanger, P. Soriano, M. Gendreau, A. Ruiz, A generic and flexible simulation-based analysis tool for ems management, submitted to International Joural of Production Research.

[3] C. Toregas, R. Swain, C. S. ReVelle, L. Bergman, The location of emergency service facilities, Operations Research 19 (1971) 1363–1373.

[4] R. L. Church, C. S. ReVelle, The maximal covering location problem, Papers of Regional Science Association 32 (1974) 101–118.

[5] M. S. Daskin, Application of an expected covering model to emergency medical service design, Decision Sciences 13 (1982) 416–439.

[6] M. S. Daskin, A maximum expected location problem : Formulation, properties and heuristic solution, Transportation Science 17 (1983) 416–439.

[7] C. S. ReVelle, K. Hogan, The maximum availability location problem, Transportation Science 23 (1989) 192–200.

[8] M. Gendreau, G. Laporte, F. Semet, Solving an ambulance location model by tabu search, Location Science 5 (1997) 75–88.

[9] C. S. ReVelle, Review, extension and prediction in emergency services siting models, European Journal of Operational Research 40 (1989) 58–69.

[10] V. Marianov, C. S. ReVelle, Siting emergency services, in: Z. Drezner (Ed.), Facility Location. A survey of Applications and Methods, New York : Springer, 1995.

[11] L. Brotcorne, G. Laporte, F. Semet, Ambulance location and relocation models, European Journal of Operational Research 147 (2003) 451–463.

[12] J. Goldberg, Operations research models for the deployment of emergency services vehicle, EMS Management Journal 1 (2004) 20–39.

[13] A. Başar, B. Çatay, T. Ünlüyurt, A taxonomy for emergency service station location problem, Operations Research Letters 6 (2012) 1147–1160.

[14] V. Bélanger, A. Ruiz, P. Soriano, Déploiement et redéploiement des véhicules ambulanciers dans la gestion des services préhospitaliers d'urgence, INFOR 50 (2012) 1–30.

[15] L. Aboueljinane, E. Sahin, Z. Jemai, A review of simulation models applied to emergency medical service operations, Computers & Industrial Engineering 66 (2013) 734–750.

[16] J. F. Repede, J. J. Bernardo, Developping and validating a decision support system for location emergency medical vehicles in Louisville, Kentucky, European Journal of Operational Research 75 (1994) 567–581.

[17] G. Carpentier, La conception et la gestion d'un réseau de service ambulancier, Master's thesis, Université Laval (2006).

[18] H. K. Rajagopalan, C. Saydam, J. Xiao, A multiperiod set covering location model for dynamic redeployment of ambulances, Computers & Operations Research 35 (2008) 814–826.

[19] A. Başar, B. Çatay, T. Ünlüyurt, A multi-period double coverage approach for locating the emergency medical service stations in Istanbul, Journal of the Operational Research Society 62 (2011) 627–637.

[20] V. Schmid, K. F. Doerner, Ambulance location and relocation problems with time-dependent travel times, European Journal of Operational Research 207 (2010) 1293–1303.

[21] C. Saydam, H. K. Rajagopalan, E. Sharer, K. Lawrimore-Belanger, The dynamic redeployment coverage location model, Health Systems (2013) 1–17.

[22] M. Gendreau, G. Laporte, F. Semet, A dynamic model and parallel tabu search heuristic for real-time ambulance relocation, Parallel Computing 27 (2001) 1641–1653.

[23] M. Gendreau, G. Laporte, F. Semet, The maximal expected relocation problem for emergency vehicles, Journal of the Operational Research Society 57 (2006) 22–28.

[24] T. Andersson, P. Värbrand, Decision support tools for ambulance dispatch and relocation, Journal of the Operational Research Society 58 (2007) 195–201.

[25] M. S. Maxwell, M. Restepo, S. G. Henderson, H. Topaloglu, Approximate dynamic programming for ambulance redeployment, INFORMS Journal on Computing.

[26] V. Schmid, Solving the dynamic ambulance relocation problem and dispatching problem using approximate dynamic programming, European Journal of Operational Research 219 (2012) 611–621.

[27] J. Naoum-Sawaya, S. Elhedhli, A stochastic optimization model for real-time ambulance redeployment, Computers & Operations Research 40 (2013) 1972–1978.

[28] A. J. Mason, Simulation and real-time optimised relocation for improving ambulance operations, in: B. Benson (Ed.), Handbook of Healthcare Operations: Methods and Applications, International Series in Operations Research & Management Science, Springer, 2013, Ch. 11, pp. 289–317.

[29] M. O. Ball, L. F. Lin, A reliability model applied to emergency service vehicle location, Operations Research 41 (1993) 18–36.

[30] K. F. Doerner, W. J. Gutjahr, R. F. Hartl, M. Karall, M. Reimann, Heuristic solution of an extended double-coverage ambulance location problem for austria, Central European Journal of Operations Research 13 (2005) 325–340.

[31] G. Laporte, F. V. Louveaux, F. Semet, A. Thirion, Applications of the double standard model for ambulance location, springer Edition, Vol. Innovations in Distribution Logistics of Lecture Notes in Economics and Mathematical System, Berlin : Springer, Berlin, 2009, pp. 235–249.

[32] A. M. Law, Simulation Modeling & Analysis, 4th Edition, Boston : McGraw-Hill, 2006.

[33] M. Pidd, Computer simulation in management science, John Wiley & Sons, 2006.

[34] A. Ingolfsson, E. Erkut, S. Budge, Simulation of single start station for Edmonton EMS, Journal of the Operational Research Society 54 (2003) 736–746.

[35] Urgences-santé, Rapport annuel, Tech. rep. (2006).
URL `http://www.urgences-sante.qc.ca/`

[36] Statistique Canada, 2011 census profile, Tech. Rep. Catalogue no. 98-316-XWE (2011).
URL `http://www12.statcan.gc.ca/census-recensement/2011/`

[37] A. Ingolfsson, EMS planning and management, in: Operations Research and Healthcare policy, International Series in Operations Research & Management Science, Springer, 2013, Ch. 6, pp. 105–128.